JOURNAL OF MULTIVARIATE ANALYSIS 8, 396–404 (1978)

Nonnull Distributions of Some Statistics Associated with Testing for the Equality of Two Covariance Matrices

B. N. NAGARSENKER

Department of Mathematics and Statistics, University of Pittsburgh

Communicated by A. T. James

The nonnull distribution of some statistics, used for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ are obtained as mixtures of incomplete beta functions as well as mixtures of incomplete gamma functions. The introduction of the convergence factors and certain recurrence relations are useful in the computation of the power of the tests as well as computation of exact percentage points for tests of significance.

1. INTRODUCTION

Let $\mathbf{X}_1 : p \times n_1$ and $\mathbf{X}_2 : p \times n_2$, $p \leqslant n_i$ $(i = 1, 2)$ be independent matrix variates with the columns of $\mathbf{X}_1$ distributed independently as $N(0, \mathbf{\Sigma}_1)$ and those of $\mathbf{X}_2$ independently distributed as $N(0, \mathbf{\Sigma}_2)$. Thus $\mathbf{S}_1 = \mathbf{X}_1\mathbf{X}_1'$ and $\mathbf{S}_2 = \mathbf{X}_2\mathbf{X}_2'$ are independently distributed as Wishart $(n_i, p, \mathbf{\Sigma}_i)$, $i = 1, 2$. Let $0 < f_1 \leqslant f_2 \leqslant \cdots \leqslant f_p < \infty$ be the characteristic roots of $\mathbf{S}_1\mathbf{S}_2^{-1}$ and $0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_p < \infty$ be those of $\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}$. Pillai and Nagarsenker [8] considered the noncentral distributions of statistics of the form

$$Y = \prod_{i=1}^{p} \theta_i^{a}(1 - \theta_i)^{b}, \tag{1.1}$$

where

$$\theta_i = f_i/(1 + f_i), \qquad i = 1, 2, \dots, p. \tag{1.2}$$

By specializing to the case where (i) $a = n_1/2$, $b = n_2/2$, (ii) $a = 1$, $b = 0$, (iii) $a = 0$, $b = 1$, one can generate various statistics used for testing the hypothesis H_0: $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$. Asymptotic expansions of the nonnull distributions of Y for the above cases have been studied by Pillai and Nagarsenker [8], Subrahmaniam [9],

Received January 1976; revised April 5, 1978.

AMS 1970 subject classifications: Primary 62H10; Secondary 62E15.

Key words and phrases: Nonnull distributions, hypergeometric functions of matrix argument, zonal polynomial, mixtures of beta functions and gamma functions.

0047-259X/78/0083–0396$02.00/0

and Khatri and Srivastava [5] for the case (i) while the exact noncentral distributions of Y for the above special cases for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ was obtained by Pillai and Nagarsenker [8] and Khatri and Srivastava [4] for the case (i). But since these density functions are given in terms of H-functions, which are just the Mellin–Barnes-type integrals, the density functions of Y obtained by these authors are in the form of an integral and so is not useful for practical purposes as far as computation of the power of the tests and the percentage points for tests of significance are concerned. Lee *et al.* [6] approximated certain powers of $\prod_{i=1}^{p} \theta_i$ with Pearson Type I distribution when $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ and found the approximation to be quite good. Tretter and Walster [10] obtained central and noncentral (linear case only) distributions of Wilks' statistic in Manova as mixtures of incomplete beta functions. In this paper the null and non-null distribution of Y for the three special cases stated above, for testing $H_0 : \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ are obtained as mixtures of incomplete beta functions and also as mixtures of incomplete gamma functions. The introduction of the adjustable constants to govern the rate of convergence of the resulting mixture representations and the recurrence relations (2.9) and (2.12) are distinctly well suited to computations of powers as well as percentage points.

2. Nonnull Distribution of Y, $a = n_1/2$ and $b = b_2/2$.

We consider below the nonnull distribution of Y first as a mixture of incomplete beta functions and then as a mixture of incomplete gamma functions.

(a) *Nonnull Distribution of Y as a Mixture of Incomplete Beta Functions*

Let

$$W = [n^n/n_1^{n_1}n_2^{n_2}]^{(1/2)p} \cdot Y. \tag{2.1}$$

Then from (2.6) of Pillai and Nagarsenker [8], the hth moment of W is given by

$$E(W^h) = C(p, n, \Lambda) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \cdot C_1(h), \tag{2.2}$$

where $(\cdot)\kappa$ and $C_\kappa(\cdot)$ are defined by James [2] and $C_1(h) = [n^n/n_1^{n_1}n_2^{n_2}]^{(1/2)ph}$.

$$\prod_{j=1}^{p} \left\{ \frac{\Gamma(\frac{1}{2}n_1(1+h) + k_j - \frac{1}{2}(j-1))\, \Gamma(\frac{1}{2}n_2(1+h) - \frac{1}{2}(j-1))}{\Gamma(\frac{1}{2}n(1+h) + k_j - \frac{1}{2}(j-1))} \right\}.$$

Using inverse Lellin transform, the density of W is given by

$$f(W) = C(p, n, \Lambda) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \cdot p(W), \tag{2.3}$$

where

$$p(W) = (2\pi i)^{-1} \int_{c-i\infty}^{c+i\infty} W^{-h-1} C_1(h)\, dh. \tag{2.4}$$

It is noted that (2.3) is valid if the condition $\max_i | 1 - Ch_i \mathbf{\Lambda} | < 1$ or $Ch_{\min}(\mathbf{\Sigma}_1 \mathbf{\Sigma}_2^{-1}) > \frac{1}{2}$ is satisfied (see [4]). Putting $1 + h = s - \lambda$ in (2.4) where λ is an adjustable constant which can be chosen to govern the rate of convergence, we have

$$p(W) = (2\pi i)^{-1} \int_{d-i\infty}^{d+i\infty} W^{-s+\lambda} C_2(s)\, ds, \tag{2.5}$$

where on using the following expansion for the gamma function,

$$\log \Gamma(x+h) = \tfrac{1}{2}\log(2\pi) + (x + h - \tfrac{1}{2}) \log x - x - \sum_{r=1}^{m} \frac{(-1)^r B_{r+1}(h)}{r(r+1)x^r} + R_{m+1}(x), \tag{2.6}$$

where $R_m(x)$ is the remainder such that $| R_m(x)| \leqslant \theta / | x |^m$, θ is a constant independent of x and $B_r(h)$ the Bernoulli polynomial of degree r and order one, $C_2(s)$ can be written in the form

$$C_2(s) = K(n) \cdot \phi(s),$$

where

$$K(n) = [(n_1/2)^{(n_1 p/2) - v + k} (n_2/2)^{(n_2 p/2) - v} / (n/2)^{(np/2) - v + k}] (2\pi)^{(1/2)p}, \tag{2.7}$$
$$v = p(p+1)/4,$$

and

$$\phi(s) = s^{-v} \left(1 + \sum_{r=1}^{\infty} \frac{l_r}{s^r}\right), \tag{2.8}$$

where the coefficients l_r can be recursively computed using the recurrence relation

$$l_r = \frac{1}{r} \sum_{k=1}^{r} k q_k l_{r-k}, \qquad l_0 = 1. \tag{2.9}$$

where

$$q_r = (-1)^{r-1} \sum_{j=1}^{p} \left[\frac{B_{r+1}(k_j - \frac{1}{2}(n_1 \lambda + (j-1)))}{(n_1/2)^r} + \frac{B_{r+1}(-\frac{1}{2}(n_2 \lambda + j - 1))}{(n_2/2)^r} - \frac{B_{r+1}(k_j - \frac{1}{2}(n\lambda + (j-1)))}{(n/2)^r} \right] \Big/ r(r+1). \tag{2.10}$$

Since

$$\phi(s) = \mathbf{O}(s^{-v}),$$

using the theorem in Nair [7], we can expand $\phi(s)$ in the factorial series as

$$\phi(s) = s^{-v}\left(1 + \sum_{r=1}^{\infty} \frac{l_r}{s^r}\right) = \sum_{i=0}^{\infty} \frac{R_i \Gamma(s+a)}{\Gamma(s+v+a+i)}, \tag{2.11}$$

where a is any arbitrary adjustable constant which can be chosen to govern the rate of convergence of the resulting series. It is easy to check the following explicit relations to determine the coefficients R_i·

$$\sum_{j=0}^{i} R_{i-j} C_{i-jj} = l_i \qquad (i = 1, 2, \ldots), \qquad R_0 = 1, \tag{2.12}$$

where

$$C_{ir} = \frac{1}{r} \sum_{k=1}^{r} k A_{ik} C_{r-k}, \qquad C_{i0} = 1,$$

and

$$A_{ij} = (-1)^{j-1}[B_{j+1}(a) - B_{j+1}(v+a+i)]/j(j+1).$$

Now using (2.11) in (2.5) and noting that term by term integration is valid since a factorial series is uniformly convergent in a half-plane (see Doetch [1]), we have

$$p(W) = K(n) \sum_{i=0}^{\infty} [(R_i)\, W^{\lambda+a}(1-W)^{v+i-1}/\Gamma(v+i)]. \tag{2.13}$$

Using (2.13) in (2.3) we have the following nonnull density function of W as a mixture of incomplete beta functions.

$$P(W \leqslant w) = C(p, n, \Lambda) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \cdot K(n)$$

$$\times \sum_{i=0}^{\infty} [(R_i')\, B_w(\lambda + a + 1, v + i)], \qquad R_0' = 1, \tag{2.14}$$

where $B_w(p, q)$ is the incomplete beta function

$$B_w(p, q) = [B(p, q)]^{-1} \int_0^w x^{p-1}(1-x)^{q-1}\, dx,$$

and

$$R_i' = (R_i)\, B(\lambda + a + 1, v + i)/\Gamma(v + i).$$

Remark. Putting $\mathbf{M} = 0$ in (2.14), we see that the null distribution W for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ is the following mixture of incomplete beta functions obtained

$$P(W \leqslant w) = C(p, n, \mathbf{I}) \cdot K_1(n) \sum_{i=0}^{\infty} [(Q_i)\, B_w(\lambda + a + 1, v + i)], \quad (2.15)$$

where $K_1(n)$ is $K(n)$ with $k = 0$ and the coefficients Q_i are the coefficients R_i' with $k_i = 0$.

(b) *Nonnull Distribution of Y as a Mixture of Incomplete Gamma Functions*

Let $L = -2q \log W$, where q is an adjustable constant which can be chosen to govern the rate of convergence of the resulting gamma series and $0 < q < \infty$ and W as in (2.1). If $H(t)$ is the characteristic function of L, then

$$H(t) = C(p, n, \mathbf{\Lambda}) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \cdot G(t), \quad (2.16)$$

where using (2.6) we can write $G(t)$ as

$$G(t) = K(n) \cdot (1 - 2qit)^{-v} \left[1 + \sum_{r=1}^{\infty} \frac{C_r}{(1 - 2qit)^r}\right]$$

where

$$v = p(p + 1)/4,$$
$$C_r = \frac{1}{r} \sum_{k=1}^{r} k A_k C_{r-k}, \qquad C_0 = 1, \quad (2.17)$$

and

$$A_r = (-1)^{r-1} \sum_{j=1}^{p} \left[\frac{B_{r+1}(k_j - \frac{1}{2}(j - 1))}{(n_1/2)^r} + \frac{B_{r+1}\frac{1}{2}(1 - j)}{(n_2/2)^r}\right.$$
$$\left. - \frac{B_{r+1}(k_j - \frac{1}{2}(j - 1))}{(n/2)^r}\right]\Big/ r(r + 1).$$

Inverting the characteristic function $H(t)$, the nonnull cdf of L is

$$P(L \geqslant L_0) = C(p, n, \mathbf{\Lambda}) \sum_{k=0}^{\infty} \sum_{k} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \cdot K(n)$$
$$\times \sum_{r=0}^{\infty} (C_r)\, G_{r+v}(2q, L_0), \qquad C_0 = 1, \quad (2.18)$$

where

$$G_\alpha(\beta, x) = [\beta^\alpha \Gamma(\alpha)]^{-1} \int_x^\infty g_\alpha(\beta, x)\, dx.$$

In particular, taking $q = 1$, we see that the nonnull cdf of L for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ may be expressed as a mixture of chi-square distributions.

Remark. Taking $\mathbf{\Lambda} = \mathbf{I}$ (i.e., $\mathbf{M} = 0$), the null distribution of L is the following mixture of incomplete gamma functions:

$$P(L \geqslant L_0) = C(p, n, \mathbf{I}) \cdot K_1(n) \sum_{r=0}^{\infty} (D_r)\, G_{r+v}(2q, L_0), \qquad D_0 = 1,$$

where $K_1(n)$ is as in Section 2(a) and D_r are the coefficients C_r with $k_r = 0$.

3. Nonnull Distribution of Y, $a = 1$ and $b = 0$

(a) *Nonnull Distribution of Y as a Mixture of Incomplete Beta Functions*

For this case, putting $a = 1$ and $b = 0$ in (2.6) of Pillai and Nagarsenker [8] and proceeding as in Section 2(a), the nonnull cdf of $Y = \prod_{i=1}^{p} \theta_i$ for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ is the following mixture of cincomplete beta functions:

$$P(Y \leqslant y) = C_1(p, n, \mathbf{\Lambda}) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!}$$

$$\times \sum_{i=0}^{\infty} \left[(R_i')\, B_y\left(\frac{n-\mu}{2} + a, \frac{pn_2}{2} + i\right)\right], \qquad R_0' = 1, \quad (3.1)$$

where

$$R_i' = (R_i)\, B\left(\frac{n-\mu}{2} + a, \frac{pn_2}{2} + i\right) \Big/ \Gamma\left(\frac{pn_2}{2} + i\right), \quad (3.2)$$

a is an adjustable constant as in Section 2(a), and the coefficients R_i are determined using the recurrence relations similar to (2.9) and (2.12).

Remark. Putting $\mathbf{M} = 0$, the null density of $Y = \prod_{i=1}^{p} \theta_i$ for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ is the following mixture:

$$P(Y \leqslant y) = C_1(p, n, \mathbf{I}) \cdot \sum_{i=0}^{\infty} \left[(D_i)\, B_y\left(\frac{n-\mu}{2} + a, \frac{pn_2}{2} + i\right)\right], \quad (3.3)$$

where the coefficients D_i are the coefficients R_i' in (3.2) with $k_i = 0$.

(b) *Nonnull cdf of* $Y = \prod_{i=1}^{p} \theta_i$ *as a Mixture of Incomplete Gamma Functions*

Let $Y_1 = -nq \log Y$, where q is the adjustable constant as in Section 2(b). Then it is easy to check that the nonnull cdf of $Y_1 = -nq \log Y$ is the following mixture of incomplete gamma functions:

$$P(Y_1 \geqslant y_0) = C_1(p, n, \Lambda) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa C_\kappa(\mathbf{M})}{k!} \sum_{r=0}^{\infty} (A_r) \left(\frac{n}{2}\right)^{-v} G_{r+v}(2q, y_0), \quad (3.4)$$

where $v = (pn_2/2)$ and the coefficients A_r can be recursively computed using recurrence relations similar to (2.17) of Section 2(b).

Remark. Putting $\mathbf{M} = 0$, the null distribution of $Y_1 = -nq \log Y$ is the following mixture:

$$P(Y_1 \geqslant y_0) = C_1(p, n, \mathbf{I}) \cdot \sum_{r=0}^{\infty} (A_r) \left(\frac{n}{2}\right)^{-v} G_{r+v}(2q, y_0),$$

where the coefficients A_r' are the coefficients A_r with $k_i = 0$.

4. Nonnull Distribution of Y, $a = 0$ and $b = 1$

(a) *Nonnull cdf of* $Y' = \prod_{i=1}^{p} (1 - \theta_i)$ *as a Mixture of Incomplete Beta Functions*

Putting $a = 0$ and $b = 1$ in (2.6) of Pillai and Nagarsenker [8] we have, proceeding as in Section 2(a), the nonnull cdf of Y' as the following mixture of incomplete beta functions:

$$P(Y' \leqslant y) = C_2(p, n, \Lambda) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa (n_1/2)_\kappa C_\kappa(\mathbf{M})}{k!} \times \sum_{i=0}^{\infty} \left[(R_i') \, B_y \left(\frac{n-\mu}{2} + a, v + i\right)\right], \quad (4.1)$$

where

$$\begin{aligned} C_2(p, n, \Lambda) &= [\Gamma_p(n/2)/\Gamma_p(n_2/2)] \cdot |\Lambda|^{-n_1/2}, \\ R_i' &= (R_i) \, B\left(\frac{n-\mu}{2} + a, v + i\right) \Big/ \Gamma(v + i), \end{aligned} \quad (4.2)$$

and

$$v = \left(\frac{pn_1}{2} + k\right).$$

a is the adjustable constant as in Section 2(a), and the coefficients R_i satisfy the recurrence relations similar to (2.9) and (2.12).

Remark. Putting $\mathbf{M} = 0$, the null density of $Y' = \prod_{i=1}^{p} (1 - \theta_i)$ for testing $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$ is the following mixture:

$$P(Y' \leqslant y) = C_2(p, n, \mathbf{I}) \sum_{i=0}^{\infty} \left[(Q_i) \, B_y \left(\frac{n - \mu}{2} + a, v + i \right) \right], \quad (4.3)$$

where the coefficients Q_i are the coefficients R_i' in (4.2) with $k_i = 0$.

(b) *Nonnull cdf of $Y' = \prod_{i=1}^{p}(1 - \theta_i)$ as a Mixture of Incomplete Gamma Functions*

Let $Y_2 = -nq \log Y'$,fwhere q is the adjustable constant as in Section 2(b). Then the nonnull cdf of $Y_2 = -nq \log Y'$ is the following mixture of incomplete gamma functions:

$$P(Y_2 \geqslant y) = C_2(p, n, \mathbf{\Lambda}) \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(n/2)_\kappa (n_1/2)_\kappa C_\kappa(\mathbf{M})}{k!}$$

$$\times \sum_{r=0}^{\infty} (A_r) \left(\frac{n}{2}\right)^{-v} G_{r+v}(2q, y), \quad (4.4)$$

where $v = (pn_1/2 + k)$ and the coefficients A_r can be recursively computed using the recurrence relations similar to (2.17).

Remark. Putting $\mathbf{M} = 0$, the null distribution of $Y_2 = -nq \log Y$ is the following mixture of incomplete gamma functions:

$$P(Y_2 \geqslant y) = C_2(p, n, \mathbf{I}) \sum_{r=0}^{\infty} (A_r') \left(\frac{n}{2}\right)^{-v} G_{r+v}(2q, y),$$

where the coefficients A_r' are the coefficients A_r in (4.4) with $k_r = 0$ and $v = pn_1/2$.

References

[1] Doetsch, G. (1971). *Guide to the Applications of the Laplace and Z-Transformations.* Van Nostrand–Reinhold, New York.

[2] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

[3] Kalinin, V. M., and Shalaevskii, O. V. (1971). *Investigations in Classical Problems of Probability Theory and Mathematical Statistics*, Part I. Consultants Bureau, New York.

[4] Khatri, C. G., and Srivastava, M. S. (1971). On exact nonnull distributions of likelihood ratio criteria for sphericity test and equality of two covariance matrices. *Sankhya Ser. A.* **33** 201–206.

[5] Khatri, C. G., and Srivastava, M. S. (1974). Asymptotic expansions of the non-null distributions of likelihobd ratio criteria for covariance matrices. *Ann. Statist.* **2** 109–117.

[6] LEE, J. C., KRISHNAIAH, P. R., AND CHANG, T. C. (1977). Approximations to the distributions of the determinants of real and complex multivariate beta matrices. *S. African Statist. J.* **11** 13–26.

[7] NAIR, U. S. (1940). Application of factorial series in the study of distribution laws in statistics. *Sankhya* **5** 175.

[8] PILLAI, K. C. S., AND NAGARSENKER, B. N. (1972). On the distributions of a class of statistics in multivariate analysis. *J. Multivariate Anal.* **2** 96–114.

[9] SUBRAHMANIAM, K. (1975). On the asymptotic distributions of some statistics used for testing $\mathbf{\Sigma} = \mathbf{\Sigma}$. *Ann. Statist.* 3 916–925.

[10] TRETTER, M. J., AND WALSTER, G. W. (1975). Central and noncentral distributions of Wilks' statistic in MANOVA as mixtures of incomplete beta functions. *Ann. Statist.* 3 467–472.